

Popularnonaukowe streszczenie projektu

Przetwarzanie dźwięku przestrzennego najczęściej kojarzone jest z symulacją przestrzeni w odtwarzanym materiale audio. Przeważająca część aplikacji związana jest z rozrywką - np. przemysłem filmowym, czy w kreowaniu wirtualnej rzeczywistości. Niemniej jednak należy pamiętać, że przetwarzanie przestrzenne stanowi również ważny aspekt w komunikacji, zwłaszcza w zastosowaniach konferencyjnych, gdzie przestrzenne rozmieszczenie poszczególnych mówców ma duży wpływ na zrozumiałość przekazu.

Rozwój metod analizy audio na przestrzeni lat umożliwił automatyzację w wielu obszarach takich jak rozpoznawanie mowy, czy identyfikacja i lokalizacja źródeł dźwięku. Niemniej jednak te metody najczęściej były rozwijane w oparciu o pojedyncze sygnały nagrywane w polu bliskim.

W projekcie zostaną zbadane możliwości integracji metod analizy sceny akustycznej (przy pomocy nagrań z wielu mikrofonów) oraz rozpoznawaniu mówcy i klasyfikacji źródeł dźwięku z wykorzystaniem uczenia maszynowego. Celem prac badawczych jest opracowanie nowych metod, w których zadania te nie będą logicznie odseparowane, ale poprzez połączenie będą mogły zostać wykorzystane jako ujednoczona metoda separacji, lokalizacji i rozpoznawania dźwięków.

W początkowych badaniach, dla każdego z powyższych obszarów nacisk zostanie położony na opracowanie sposobów łączenia w celu poprawy skuteczności każdego z nich. Analiza sceny akustycznej nagranej przy pomocy macierzy mikrofonowej pozwoli uzyskać cechy wspomagające techniki uczenia maszynowego, w szczególności klasyfikację źródeł dźwięku oraz rozpoznawanie mówcy. Dotychczas głównym celem badań było zminimalizowanie wpływu przestrzeni akustycznej na metody klasyfikacji źródeł. W tym projekcie badania będą skupiały się na odkryciu i wykorzystaniu informacji pochodzących z analizy tej przestrzeni do polepszenia skuteczności rozpoznania. Analogicznie informacje pochodzące z klasyfikatorów mogą poprawić skuteczność algorytmów przetwarzania przestrzennego dźwięku np lokalizacji źródeł. Spodziewanym rezultatem projektu jest stworzenie algorytmu, który jednocześnie identyfikuje i klasyfikuje źródło dźwięku z większą skutecznością niż przy użyciu modułów analizy przestrzeni dźwięku, rozpoznawania mówcy oraz klasyfikacji źródeł dźwięku oddzielnie. Pośrednim efektem projektu będzie więc nie tylko rozwój w poszczególnych obszarach przetwarzania audio, ale również zbadanie synergii między nimi.